Tree Biotechnology 2015 Conference

8-12 June 2015
Florence, Italy

IUFRO Tree Biotechnology 2015 Conference

"Forests: the importance to the planet and society"

8-12 June 2015
Florence, Italy

Proceedings

Editors

Cristina Vettori
Institute of Bioscience and BioResouces
National Research Council
Sesto Fiorentino (FI)
Italy

Giovanni Giuseppe Vendramin
Institute of Bioscience and BioResouces
National Research Council
Sesto Fiorentino (FI)
Italy

Donatella Paffetti
Department of Agriculture, Food and Environmental Sciences
University of Florence
Florence
Italy

Davide Travaglini
Department of Agricultural, Food and Forestry Systems
University of Florence
Florence
Italy

DOI 10.13140/RG.2.1.4603.6882

Cite these articles as in this example:

Vettori et al. 2015. Gene expression profiling in response to UV-B radiation in different Populus alba clones. *In Proceedings of the IUFRO Tree Biotechnology 2015 Conference: "Forests: the importance to the planet and society"*. Eds: Vettori C, Vendramin G.G., Paffetti D., Travaglini D. - doi: 10.13140/RG.2.1.4603.6882; ID: S1.P2

Proceedings of the IUFRO Tree Biotechnology 2015 Conference
Florence, Italy – 8/12 June, 2015
Poster Communication Abstract – ID: S3.P46
DOI 10.13140/RG.2.1.4603.6882

De novo assembly and annotation of a transcriptome during xylogenesis in *Pinus canariensis* Chr. Sm. ex DC

Víctor Chano[1,a], Unai López de Heredia[1,b], Carmen Collada[1,c], Álvaro Soto[1,d]

[1]GENFOR, Grupo de Investigación en Genética y Fisiología Forestal. ETSI Montes. Universidad Politécnica de Madrid. Ciudad Universitaria s/n 28040 Madrid, Spain. [a]vm.chano@upm.es; [b]unai.lopezdeheredia@upm.es; [c]carmen.collada@upm.es; [d]alvaro.soto.deviana@upm.es.

Key words: Pinus canariensis, Transcriptome, Annotation, Next-generation sequencing

Background

Nowadays, there is a great amount of genomic and transcriptomic data available about forest species, including ambitious projects looking for complete sequencing and annotation of different gymnosperm genomes [1, 2]. *Pinus canariensis* is an endemic conifer of the Canary Islands with re-sprouting capability and resilience against fire and mechanical damage, as result of an adaptation to volcanic environments. Additionally, this species has a high proportion of axial parenchyma compared with other conifers, and this tissue connects with radial parenchyma allowing transport of reserves. The most internal tracheids stop accumulating water [3], and get filled of resins and polyphenols synthesized by the axial parenchyma; this is the so-called “torch-heartwood” [4], which avoids decay. This wood achieves very high prices due to its particular resistance to rot. These features make *P. canariensis* an interesting model species for the analysis of these developmental processes in conifers. In this study we aim to perform a complete transcriptome annotation during xylogenesis in *Pinus canariensis*, using next-generation sequencing (NGS) -Roche 454 pyrosequencing-, in order to provide a genomic resource for further analysis, including expression profiling and the identification of candidate genes for important adaptive features.

Methods

Differentiating xylem was collected at two points of seasonal growth, spring and summer, in order to cover all anatomical events. Total RNA was extracted based on Chang *et al.* [5], and two separate libraries were constructed through 454 pyrosequencing and *de novo* assembled using Newbler 2.5 (Roche454; [6]). Preliminary assemblies were pooled and meta-assembled using CAP3 [7], to remove redundancies and achieve larger contigs. The quality of the assembling was assessed quantitavely by computing the length of contigs, the GC% content and the N50 using Quast 2.3 [8]. Contigs were launched in a local version of BlastX [9] against the Viridiplantae section of RefSeq database (NCBI), with a threshold *evalue* of 0.00001. BlastX output was imported into Blast2GO [10] to assign Gene Ontology (GO) information. Contigs were also aligned to available transcriptomes of *Pinus pinaster* and *Pinus halepensis* [1, 11].

Results and Conclusions

The two libraries constructed at two stages of vegetative growth of *P. canariensis* bring 458,498 reads in spring and 474,393 reads in summer. After assembly, the final transcriptome resulted in 45,509 contigs and 103,764 singletons (i.e. short reads not asembled into contigs), having a total length of 34,024,623 bp, and with a GC% of 42.4 and N50 size of 1,078 bp. BlastX search gave 27,474 hits (60.37 % of contigs), and 172,687 GO terms were obtained after annotation, distributed among different levels for Biological Process (BP; 93,157 GOs, 54%), Molecular Function (MF; 32,994 GOs, 19%), and Cellular Component (CC; 46,536 GOs, 27%). These data were similar to those obtained in *P. halepensis* through RNAseq [11] and Douglas fir with 454 methods [12]. Finally, 92.66% and 79.27% of the contigs aligned to unigenes of *P. pinaster* and *P. halepensis*, respectively.

This work is the first comprehensive analysis of the *P. canariensis* transcriptome, and our data may be a useful resource for further analysis in the identification of candidate genes for a better understanding of adaptive patterns in conifers, such as cell-types differentiation, re-sprout, wound-healing or “torch-heartwood” formation.

Competing interests
The author declares that they have no competing interests.

References

1. Fernández-Pozo N, Canales J, Guerrero-Fernández D, Villalobos DP, Díaz-Moreno SM, Bautista R, Flores-Monterroso A, Guevara MÁ, Perdiguero P, Collada C, Cervera MT, Soto A, Ordás R, Cantón FR, Avila C, Cánovas FM, Claros MG: **EuroPineDB: a high-coverage web database for maritime pine transcriptome.** *BMC Genomics* 2011, **12**:366.

2. Zimin A, Stevens K a, Crepeau MW, Holtz-Morris A, Koriabine M, Marçais G, Puiu D, Roberts M, Wegrzyn JL, de Jong PJ, Neale DB, Salzberg SL, Yorke J a, Langley CH: **Sequencing and assembly of the 22-gb loblolly pine genome.** *Genetics* 2014, **196**:875–90.

3. Climent J, Chambel MR, Pérez E, Gil L, Pardos J: **Relationship between heartwood radius and early radial growth , tree age , and climate in *Pinus canariensis***. *Can J For Res* 2002, **32**:103–111.

4. Climent J, Gil L, Pardos JA: **Xylem anatomical traits related to resinous heartwood formation in *Pinus canariensis* Sm .** *Trees* 1998, **12**:139–145.

5. Chang S, Puryear J, Cairney J: **A simple and efficient method for isolating RNA from pine trees**. *Plant Mol Biol Report* 1993, **11**:113–116.

6. Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes X V, Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, *et al*.: **Genome sequencing in microfabricated high-density picolitre reactors.** *Nature* 2005, **437**:376–380.

7. Huang X, Madan A: **CAP3: A DNA sequence assembly program**. *Genome Res* 1999, **9**:868–877.

8. Gurevich A, Saveliev V, Vyahhi N, Tesler G: **QUAST: Quality assessment tool for genome assemblies**. *Bioinformatics* 2013, **29**:1072–1075.

9. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ: **Basic local alignment search tool.** *J Mol Biol* 1990, **215**:403–410.

10. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M: **Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research.** *Bioinformatics* 2005, **21**:3674–6.

11. Pinosio S, González-Martínez SC, Bagnoli F, Cattonaro F, Grivet D, Marroni F, Lorenzo Z, Pausas JG, Verdú M, Vendramin GG: **First insights into the transcriptome and**

development of new genomic tools of a widespread circum-Mediterranean tree species, *Pinus halepensis* Mill. *Mol Ecol Resour* 2014, **14**:846–856.

12. Howe GT, Yu J, Knaus B, Cronn R, Kolpak S, Dolan P, Lorenz WW, Dean JFD: **A SNP resource for Douglas-fir: de novo transcriptome assembly and SNP detection and validation.** *BMC Genomics* 2013, **14**:137.